



# La statistica ...che cosa è? Perché viene utilizzata?

È una disciplina che concerne **la raccolta, l'analisi, l'interpretazione e presentazione dei dati.**

Nasce dalla necessità di trarre conclusioni e informazioni da un insieme di dati che costituisce solo una parte dell'intera popolazione.

## DEFINIZIONE di STATISTICA:

**Analisi quantitativa di fenomeni collettivi**, allo scopo di descriverli e di individuare leggi e modelli che classificando le loro variazioni permettono di spiegarli e di prevederli.

# Scopi e funzioni

Rendere utilizzabili grandi quantità di informazioni, teoricamente disponibili, ma di fatto difficilmente gestibili.



FUNZIONE DESCRITTIVA

FUNZIONE PREDITTIVA

Tramite la statistica è possibile sostituire ad un'impressione qualitativa di un fenomeno la sua misura quantitativa.

# Scopi e funzioni

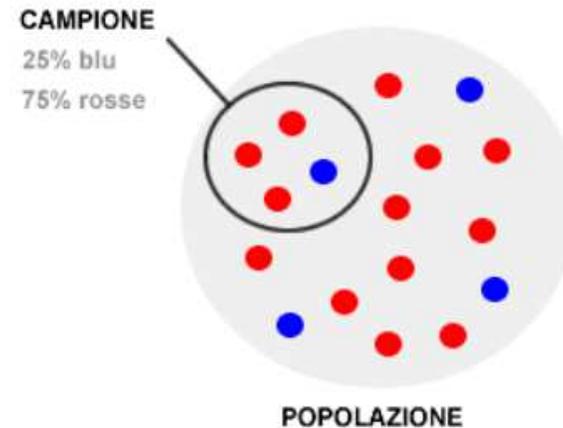
TIPI DI STATISTICA:

## Statistica DESCRITTIVA:

procedure per riassumere e presentare i dati e per descriverli attraverso strumenti matematici

## Statistica INFERENZIALE:

procedure per analizzare i dati ottenuti da un campione della popolazione per stimare un fenomeno statistico sull'intera popolazione di riferimento



# La statistica può descrivere:

TUTTA LA POPOLAZIONE (STATISTICA DESCRITTIVA)

– La metodologia impiegata è puramente descrittiva ed è un insieme di tecniche relative a questo tipo di indagini.

UN CAMPIONE DELLA POPOLAZIONE (STATISTICA INFERENZIALE)

– La metodologia impiegata può essere detta inferenziale.

Nella statistica inferenziale si devono compiere tutte le operazioni di schematizzazione e di descrizione proprie della statistica descrittiva; ad essi poi vanno aggiunti i procedimenti inferenziali

# La statistica descrittiva:

metodi per organizzare, riassumere e presentare i dati in modo informativo.

## **Esempio 1:**

Il 35% degli intervistati conosce le modalità di trasmissione del virus HIV.

La statistica 35 descrive il numero di persone che conosce le modalità di trasmissione del virus HIV tra 100 intervistate.

## **Esempio 2:**

In accordo ad interviste ai consumatori, i cellulari di marchio X hanno avuto 5 guasti per 100 durante il 2020.

La statistica 5 descrive il numero di guasti su 100 cellulari esaminati.

# La statistica inferenziale:

procedure per lo studio dei dati campionari che consentono di valutare la possibilità di generalizzazione dei risultati descrittivi, ottenuti dall'elaborazione dei dati campionari, alle popolazioni da cui sono stati estratti.

## **Esempio 1:**

La direzione sanitaria di una azienda ospedaliera vuole selezionare un campione di cartelle cliniche per verificare l'accuratezza delle diagnosi in rapporto ai drg (diagnosis related groups).

# Statistica descrittiva

Il principale obiettivo della statistica descrittiva è la conoscenza quantitativa dei fenomeni collettivi.

- I fenomeni che la nostra mente non può conoscere con una sola osservazione, ma che invece apprende tramite la sintesi delle osservazioni di fenomeni più semplici, vengono abitualmente chiamati *fenomeni collettivi*.

## **ESEMPI**

*Disoccupazione dei laureati di età compresa fra i 24 e i 35 anni in Italia*

*Recidiva nei reati di violenza sessuale in Italia*

*Delittuosità in Italia*

*Evasione fiscale delle piccole e medie imprese in Abruzzo*

*Terrorismo in Europa*

# DEFINIZIONI PRELIMINARI:



- Popolazione; Campione; Unità statistiche
- Carattere, modalità
- Classificazione dei caratteri

# Unità statistica

- L'oggetto dell'osservazione di ogni fenomeno individuale che costituisce il fenomeno collettivo è detto **unità statistica**.

**L'unità statistica** è l'unità elementare su cui viene osservato il fenomeno oggetto di studio.

Fenomeno	Unità statistica
<i>Disoccupazione dei laureati di età compresa fra i 24 e i 35 anni in Italia nel 2019</i>	<i>ogni laureato residente in Italia di età compresa fra i 24 e i 35 anni</i>
<i>Recidiva nei reati di violenza sessuale in Italia</i>	<i>ogni autore di reato di violenza sessuale in Italia</i>
<i>Delittuosità in Italia dal 2010 al 2020</i>	<i>ogni delitto commesso in Italia dal 2010 al 2020</i>
<i>Evasione fiscale delle piccole e medie imprese in Abruzzo nel 2019</i>	<i>ogni piccola e media impresa abruzzese attiva nel 2019</i>
<i>Terrorismo in Europa</i>	<i>ogni atto di terrorismo in Europa dal 2010 al 2020</i>

# Collettivo statistico

L'insieme delle unità statistiche uguali rispetto ad alcune loro caratteristiche costituiscono il **collettivo statistico**

- Nel caso di indagine statistica, il collettivo può comprendere tutte le unità omogenee rispetto alle caratteristiche prescelte (**POPOLAZIONE**) o solo una parte di esse (**CAMPIONE**).

Fenomeno	Unità statistica	Collettivo
<i>Disoccupazione dei laureati di età compresa fra i 24 e i 35 anni in Italia nel 2019</i>	<i>ogni laureato residente in Italia di età compresa fra i 24 e i 35 anni</i>	<i>campione di laureati residenti in Italia di età compresa fra i 24 e i 35 anni</i>
<i>Recidiva nei reati di violenza sessuale in Italia</i>	<i>ogni autore di reato di violenza sessuale in Italia</i>	<i>insieme degli autori di reati sessuali in Italia</i>
<i>Delittuosità in Italia dal 2010 al 2020</i>	<i>ogni delitto commesso in Italia dal 2010 al 2020</i>	<i>insieme dei delitti commessi in Italia dal 2010 al 2020</i>
<i>Evasione fiscale delle piccole e medie imprese in Abruzzo nel 2019</i>	<i>ogni piccola e media impresa abruzzese attiva nel 2019</i>	<i>campione di piccole e medie imprese abruzzesi attive nel 2019</i>
<i>Terrorismo in Europa</i>	<i>ogni atto di terrorismo in Europa dal 2010 al 2020</i>	<i>insieme degli atti di terrorismo commessi in Europa dal 2010 al 2020</i>

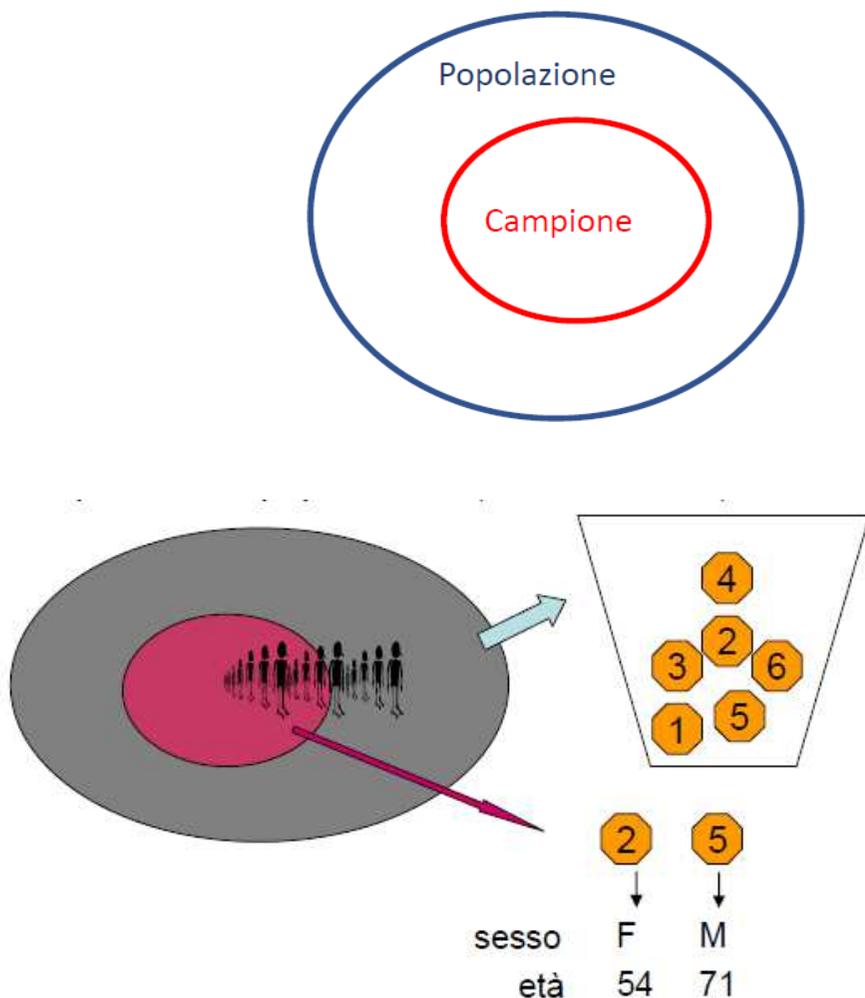
# Carattere o Variabile

Si chiama **carattere o variabile** ogni aspetto elementare oggetto di rilevazione sulle unità statistiche del collettivo.

Ogni unità statistica è portatrice, generalmente, di molteplici caratteri.

Unità statistica	Caratteri
<i>laureato residente in Italia di età compresa fra i 24 e i 35 anni</i>	<i>età, sesso, tipo di laurea conseguita, luogo di residenza, condizione occupazionale</i>
<i>autore di reato di violenza sessuale in Italia</i>	<i>età, sesso, titolo di studio, residenza, tipo di reato commesso, durata della condanna, recidiva</i>
<i>delitto commesso in Italia dal 2010 al 2020</i>	<i>Tipologia di delitto, luogo del delitto, data del delitto</i>
<i>piccola e media impresa abruzzese attiva nel 2019</i>	<i>Natura giuridica, sede legale, settore produttivo, numero di dipendenti</i>
<i>atto di terrorismo in Europa dal 2010 al 2020</i>	<i>Tipologia di reato, luogo, data, matrice</i>

# Campione



- Il campione è un sottoinsieme della popolazione.
- Viene utilizzato perché non è sempre possibile fare valutazioni sulla popolazione (cosa che fa il censimento)
- La tecnica di estrazione di un campione viene detta campionamento

# Matrice dei dati

caratteri (variabili)

unità statistiche

paziente	sexso	età	BMI	peso*	patologia	diabete
A	F	54	20.2	normopeso	diabete	sì
B	M	64	31.2	obesità	dislipidemia	no
C	M	32	17.8	sottopeso	diabete	sì
D	F	74	23.1	normopeso	insuff. renale	no
...						

# Caratteri e modalità

(modalità)

paziente	sexso	età	BMI	peso	patologia	diabete
A	F	54	20.2	normopeso	diabete	sì
B	M	64	31.2	obesità	dislipidemia	no
C	M	32	17.8	sottopeso	diabete	sì
D	F	74	23.1	normopeso	insuff. renale	no
...						

- Le caratteristiche di interesse delle unità statistiche sono dette **CARATTERI, o VARIABILI**
- I caratteri presentano (si esprimono attraverso) dei **VALORI o MODALITA'**
  - Le unità statistiche differiscono fra loro per le modalità che esse presentano: il carattere presenta una variabilità che è l'oggetto di studio della statistica.

# Modalità

Si chiamano **modalità** di un carattere i diversi modi con cui questo si presenta nelle unità statistiche del collettivo.

Carattere	Modalità
Sesso	Maschio, Femmina
Titolo di studio	Licenza elementare, Licenza media, Diploma, Laurea, Post laurea
Condizione occupazionale	Occupato, Disoccupato, Inattivo
Età	18 anni, 19 anni, ....
Regione di residenza	Abruzzo, Basilicata,....
Livello di soddisfazione	Basso, Medio, Alto

# Classificazione dei caratteri

La natura del carattere dipende da che modalità esso presenta

QUALITATIVI	QUANTITATIVI
<b>SCONNESSI</b> sesso → M,F patologia → ulcera, tumore gastrico, tumore intestinale, ...	<b>DISCRETI</b> numero di componenti (della famiglia) → 1,2,3,4, ... gravidanze precedenti → 0, 1, 2, 3, ...
<b>ORDINATI</b> titolo di studio → nessuno o licenza elementare, licenza media, licenza superiore, laurea stadio malattia → I,II,III	<b>CONTINUI</b> età (anni compiuti) → 0,1,2,...,24,...,88,... peso (kg) → 56.4, 78.2, ... WBC (x 10 <sup>3</sup> /ml) → 3.4, 2.8, ...

# Caratteri qualitativi

Presentano modalità che corrispondono a diciture, attributi, caratteristiche descrivibili attraverso “parole” (ovvero, attraverso numeri che però non corrispondono a conteggi o misurazioni, ma esprimono convenzioni)

– Non ammettono operazioni matematiche!!

- **SCONNESSI**: non si ha un ordinamento naturale o “tipico” (stabilito per convenzione)
  - è possibile solo dire se due unità sono uguali o diverse (se presentano la stessa modalità o modalità diverse)
- **ORDINATI**: esiste un ordinamento naturale o “tipico”
  - è possibile stabilire relazioni di superiorità / inferiorità fra due unità;
  - non è però possibile (o non ha senso) calcolare delle differenze per stabilire la “distanza” fra due unità
  - (Non farsi ingannare dalle codifiche numeriche!!)

# Caratteri qualitativi

## Nominali

Gruppi sanguigni: A, B, AB, 0

Fumo: Si, No, Ex fumatore

Regioni: Toscana, Lombardia, Sicilia,..

Sesso: Maschio, Femmina

Fumo: Si, No

Istologia: nodulo benigno, tumore

## Ordinali:

BMI: sottopeso, normopeso,  
sovrappeso, obeso

Esposizione al sole: scarsa, media, alta

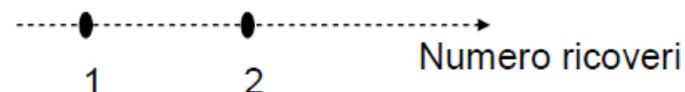
Scala NRS: 0,1,2,3,4,5,6,7,8,9,10

Indice di Apgar: 0,1,2,3,4,5,6,7,8,9,10

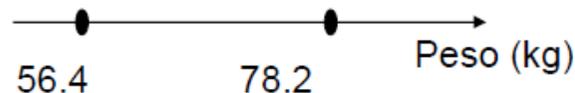
# Caratteri quantitativi

Presentano modalità effettivamente numeriche, ottenute tramite conteggio o misurazione; sulle modalità è possibile eseguire operazioni matematiche.

- **DISCRETI**: le modalità possono essere enumerate; i valori compresi fra due modalità possono NON essere a loro volta delle modalità
  - generalmente ottenuti tramite conteggio



- **CONTINUI**: le modalità NON possono essere enumerate; i valori compresi fra due modalità sono sempre a loro volta delle modalità
  - generalmente ottenuti tramite misurazione



# Caratteri quantitativi

## Discrete:

Accessi al PS: 45 72 121 3 ...

Numero di pz visitati al mese con  
sospetto melanoma: 2 45 31 12..

## Continue:

Peso: 65,7kg 80,2kg 56,1kg ...

Diametro: 2,3cm 1,7cm 2,2cm..

Altezza: 170cm 165cm 183cm..

# Operazioni sulle modalità

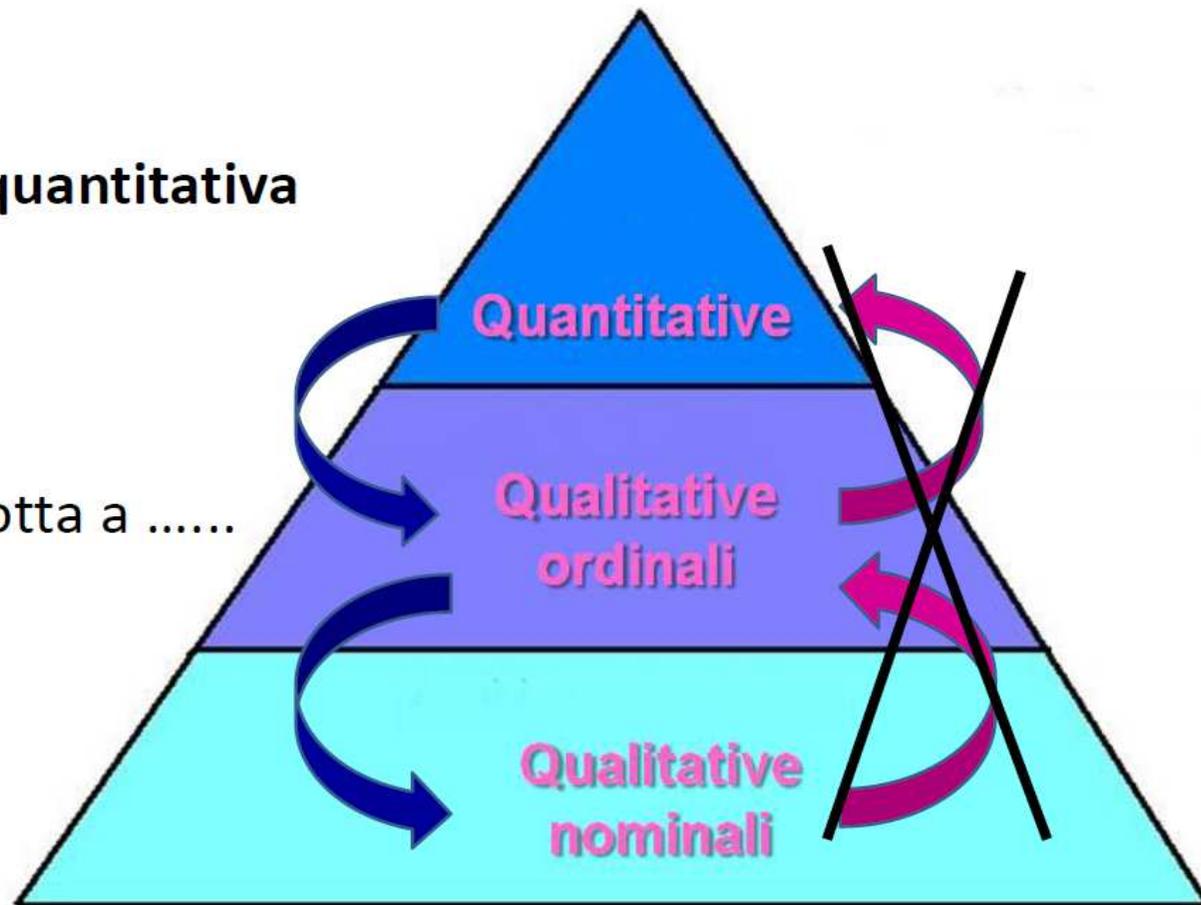
		Operazioni		
		Uguaglianza/ Disuguaglianza	Ordinamento	Operazioni aritmetiche
Caratteri		=, ≠	>, <	+, -, *, /
Qualitativi	<i>sconnessi</i>	si	no	no
	<i>ordinabili</i>	si	si	no
Quantitativi		si	si	si

# Gerarchia tra le variabili

Una **variabile quantitativa**

può essere ridotta a .....

e ancora a .....



# Gerarchia tra le variabili

1) L'*Età* dei pazienti può essere espressa a livello ordinale?



Bambini (0-10) /Adolescente (11-18) /Adulti (19-64)/Anziani (65+)

2a) L'*IMC* (*Indice di Massa Corporea, kg/m<sup>2</sup>*) può essere espressa a livello ordinale?



Sottopeso (<18)/ Normo (18-25)/ Sovra (26-29)/ Obeso (≥30)

2b) e a livello dicotomico?



Obeso (≥30) / Non Obeso (<30)

**A che tipologia appartengono i seguenti caratteri:**

1. Sesso
2. Regione di residenza
3. Titolo di studio
4. N. componenti
5. Età

id	Sesso	Regione di residenza	Titolo di studio	Numero componenti famiglia	Età
4	Femmina	Lazio	Licenza Media	2	21
5	Maschio	Puglia	Licenza Media	1	23
6	Femmina	Campania	Diploma	3	21
8	Femmina	Campania	Licenza Media	3	27
11	Maschio	Campania	Licenza Media	1	27
12	Femmina	Marche	Licenza Media	2	27
13	Femmina	Campania	Licenza Media	2	27
14	Femmina	Campania	Diploma	1	22
15	Femmina	Campania	Licenza Media	3	27
17	Femmina	Campania	Licenza Media	4	55
18	Femmina	Campania	Licenza Media	4	26
20	Femmina	Molise	Laurea	3	22
21	Maschio	Campania	Diploma	2	24
22	Femmina	Campania	Laurea	3	44
24	Femmina	Campania	Laurea	4	30
25	Femmina	Campania	Laurea	4	21
27	Maschio	Campania	Diploma	4	27
28	Femmina	Molise	Laurea	3	22
29	Femmina	Campania	Diploma	4	22
30	Femmina	Puglia	Diploma	2	62
31	Maschio	Puglia	Diploma	4	22

# DISTRIBUZIONI DI FREQUENZE



- Frequenze assolute, relative e percentuali
- Frequenze cumulate
- Tabelle doppie

# Distribuzioni statistiche

- L'effetto dell'operazione di determinazione della modalità con cui ognuno dei caratteri si presenta in ciascuna unità del collettivo determina la "**distribuzione**" del collettivo secondo i caratteri considerati.
- La distribuzione indica come le **modalità** dei **caratteri** si **distribuiscono** nelle **unità del collettivo**.

## Distribuzioni disaggregate

Sesso	Titolo di studio
Femmina	Licenza Media
Maschio	Licenza Media
Femmina	Diploma
Femmina	Licenza Media
Maschio	Licenza Media
Femmina	Laurea
Femmina	Laurea

## Distribuzioni di frequenza

modalità	frequenza
Femmina	5
Maschio	2
modalità	frequenza
Licenza Media	4
Diploma	1
Laurea	2

# Le tabelle di frequenze

- La prima operazione utile per sintetizzare una serie di dati relativa ad un carattere è il **conteggio**:

ad ogni modalità (o classe, intervallo di valori) si associa la **frequenza**, ossia il numero di unità che presentano quella modalità (o cadono in quella classe).

unità	SESSO	ETA
1	M	55
2	F	51
3	F	44
4	M	62
5	M	48
6	F	51
7	M	69
8	F	58
9	F	72
10	M	50
11	F	78
12	F	46



SESSO	n
M	5
F	7
tot	12



ETA'	n
-   50	4
50 -   65	5
65 -	3
tot	12

# Distribuzioni di frequenze

Per **frequenza** si intende il numero di volte che una data modalità si presenta nel collettivo statistico.

Si chiama **distribuzione di frequenze** lo schema con cui si associa a ciascuna modalità del carattere  $X$  la rispettiva frequenza.

<i>Modalità</i> ( $x_i$ )	<i>Frequenza</i> ( $n_i$ )
$x_1$	$n_1$
$x_2$	$n_2$
$\vdots$	$\vdots$
$x_k$	$n_k$
Totale	$N$

$$\sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k = N$$

$n_1, n_2, \dots, n_k$  sono le frequenze delle modalità  $x_1, x_2, \dots, x_k$

# Frequenze assolute, relative e percentuali

- **Assolute** : numero di volte in cui si presenta una modalità della variabile
- **Relative** : vengono calcolate come rapporto tra le frequenze assolute e la numerosità del collettivo
- **Percentuali**: vengono calcolate come prodotto delle frequenze relative per 100

# Frequenze relative e percentuali

- **frequenze relative o proporzioni**: si ottengono rapportando le frequenze assolute al totale delle unità,  $N$ . Indicheremo con  $f_1, f_2, \dots, f_k$  tali quantità, essendo

$$f_i = \frac{\text{frequenza della modalità } x_i}{\text{numero totale di osservazioni}} = \frac{n_i}{N}, i=1, 2, \dots, k.$$

- **frequenze percentuali**  $p_1, p_2, \dots, p_k$  : si ottengono moltiplicando per 100 le frequenze relative:

$$p_i = f_i * 100, \quad i=1, 2, \dots, k$$

$$\sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k = 1$$

$$\sum_{i=1}^k p_i = p_1 + p_2 + \dots + p_k = 100$$

# Frequenze relative

Variabile qualitativa «fumo»  
(0=no, 1=si, 2=ex fumatore)

Osservazioni sulle unità statistiche:  
0,1,2,0,0,0,1,2,1,2,1,0,1,2,0,1,2,0,1,1,1

	Frequenze assolute	Frequenze relative
0 (SI)	7	0.33
1 (NO)	9	0.43
2 (ex fumatore)	5	0.24
Totale	21	1

Rapporto delle frequenze  
assolute per il totale:

$$7/21=0.33$$

$$9/21=0.43$$

$$5/21=0.24$$

# Frequenze percentuali

Variabile qualitativa **fumo**  
(0=no, 1=si, 2=ex fumatore)

Osservazioni sulle unità statistiche:  
0,1,2,0,0,0,1,2,1,2,1,0,1,2,0,1,2,0,1,1,1

	Frequenze assolute	Frequenze relative	Frequenze percentuali
0 (SI)	7	0.33	33%
1 (NO)	9	0.43	43%
2 (ex fumatore)	5	0.24	24%
Totale	21	1	100%

Prodotto delle frequenze  
relative per 100%:

$$(0.33 \times 100)\% = 33\%$$

$$(0.43 \times 100)\% = 43\%$$

$$(0.24 \times 100)\% = 24\%$$

# Percentuali: interpretazione ed uso

Consideriamo i risultati di uno studio relativo

Efficacia strategia di MARKETING

Presentiamo 2 scenari in cui le frequenze percentuali di risposta sono le stesse.

EFFICACIA	freq.
No	600
Si	800
<i>tot</i>	1400

EFFICACIA	freq.
No	6
Si	8
<i>tot</i>	14

EFFICACIA	%
No	42.9
Si	57.1
<i>tot</i>	100.0

In presenza di percentuali, guardiamo e riportiamo sempre la numerosità totale del collettivo!!

## Esempio di distribuzione di frequenze assolute, relative e percentuali

**Numero di imprese attive del settore artistico, sportivo, dell'intrattenimento e del divertimento per attività economica - Anno 2020**

<b>ATTIVITÀ ECONOMICHE</b>	<b>Imprese attive</b>		
	Numero	frequenza relativa	%
Attività creative, artistiche e di intrattenimento	31,950	0.45	44.8
Attività sportive	15,455	0.22	21.7
Attività ricreative e di divertimento	14,467	0.20	20.3
Attività riguardanti le lotterie, le scommesse, le case da gioco	8,365	0.12	11.7
Attività di biblioteche, archivi, musei ed altre attività culturali	1,031	0.01	1.4
<b>Totale attività artistiche, sportive, di intrattenimento e divertimento</b>	<b>71,268</b>	<b>1</b>	<b>100</b>

Fonte Istat: STATISTICHE CULTURALI - ANNO 2021

### **Interpretazione:**

Le attività economiche inerenti il settore della cultura in Italia per l'anno 2020 si concentrano maggiormente nell'ambito delle attività creative, artistiche e di intrattenimento (45%), delle attività sportive (22%) e delle attività ricreative e di divertimento (20%).

Il peso percentuale delle attività economiche relative a biblioteche, archivi, musei a altre è molto contenuto (1%)

# Frequenze cumulate

Consideriamo una distribuzione di frequenze secondo un carattere a modalità ordinabili.

- Si chiamano **frequenze cumulate** le quantità

$$N_i = n_1 + n_2 + \dots + n_i, i = 1, 2, \dots, k.$$

Per ogni dato  $i$ ,  $N_i$  rappresenta il numero delle unità del collettivo nelle quali il carattere  $X$  assume un valore non superiore a  $x_i$

- Si chiamano **frequenze relative cumulate** i rapporti

$$F_i = \frac{\text{frequenza cumulata fino a } x_i}{\text{numero totale di osservazioni}} = \frac{N_i}{N}, i = 1, 2, \dots, k.$$

- **frequenze percentuali cumulate  $P_i$** : si ottengono moltiplicando per 100 le frequenze relative cumulate

# Esempio di frequenze cumulate

	Attese sulla situazione economica dell'Italia					
	gennaio 2023			febbraio 2023		
	%	% su valide	% cumulata	%	% su valide	% cumulata
peggiorerà molto	13.3	13.5	13.5	10.7	10.8	10.8
peggiorerà lievemente	24.2	24.6	38.1	21.7	22.0	32.8
rimarrà stabile	39.5	40.1	78.3	44	44.5	77.3
migliorerà leggermente	20.8	21.1	99.4	21.3	21.6	98.9
migliorerà molto	0.6	0.6	100	1.1	1.1	100
<i>non sa</i>	<i>1.6</i>			<i>1.2</i>		
<b>Totale</b>	<b>100</b>	<b>100</b>		<b>100</b>	<b>100</b>	

Fonte Istat: FIDUCIA DEI CONSUMATORI E DELLE IMPRESE - FEBBRAIO 2023

## **Interpretazione:**

Il 38% dei rispondenti ritiene che la situazione economica in Italia per gennaio 2023 peggiorerà molto o leggermente. Tale percentuale scende al 33% a febbraio 2023.

Il 78% dei rispondenti ritiene che la situazione economica in Italia per gennaio 2023 peggiorerà molto o leggermente o che rimarrà invariata. Tale percentuale rimane sostanzialmente invariata (77%) a febbraio 2023.

# Frequenze cumulate

	Attese sulla situazione economica dell'Italia					
	gennaio 2023			febbraio 2023		
	%	% su valide	% cumulata	%	% su valide	% cumulata
peggiorerà molto	13.3	13.5	13.5	10.7	10.8	10.8
peggiorerà lievemente	24.2	24.6	38.1	21.7	22.0	32.8
rimarrà stabile	39.5	40.1	78.3	44	44.5	77.3
migliorerà leggermente	20.8	21.1	99.4	21.3	21.6	98.9
migliorerà molto	0.6	0.6	100	1.1	1.1	100
<i>non sa</i>	1.6			1.2		
<b>Totale</b>	100	100		100	100	

Fonte Istat: FIDUCIA DEI CONSUMATORI E DELLE IMPRESE - FEBBRAIO 2023

## **Interpretazione:**

Il 38% dei rispondenti ritiene che la situazione economica in Italia per gennaio 2023 peggiorerà molto o leggermente. Tale percentuale scende al 33% a febbraio 2023.

Il 78% dei rispondenti ritiene che la situazione economica in Italia per gennaio 2023 peggiorerà molto o leggermente o che rimarrà invariata. Tale percentuale rimane sostanzialmente invariata (77%) a febbraio 2023.

# Raggruppamento in classi

Quando il carattere è quantitativo e il numero di osservazioni è elevato, la presentazione dei dati richiede necessariamente che le modalità contigue siano aggregate tramite la formazioni di classi, cioè di intervalli numerici comprendenti più modalità.

Durata della malattia	Valori assoluti	Valori %
0-2.99	114	19.8
3-5.99	188	32.6
6-8.99	140	24.3
9-11.99	87	15.1
12-14.99	43	7.5
15-17.99	5	0.9
Totale	577	

Classi Omogenee

di pari ampiezza

$$A_i = (x_{i+1} - x_i)$$

Durata della malattia	Valori assoluti	Valori %
0-2.99	114	19.8
3-7.99	293	50.8
8-10.99	93	16.1
11-13.99	65	11.3
14-15.99	10	1.7
16-17.9	2	0.3
Totale	577	

Classi Disomogenee

di ampiezza diversa

$$A_i = (x_{i+1} - x_i)$$

per  $i = 1, \dots, k$  intervalli (classi)

# Tabelle doppie

Quando per ogni unità del collettivo consideriamo due caratteri contemporaneamente otteniamo una distribuzione doppia.

Outcome/variabile dipendente

Exposure /variabile indipendente

	Mod1	Mod2	Totale di riga
Mod1			$R_1$
Mod2			$R_2$
Totale di colonna	$C_1$	$C_2$	Totale (n)

Modalità o categorie della variabile

- Frequenze assolute
- Frequenze relative
- Frequenze percentuali

Numerosità campionaria

# Tabelle doppie: frequenze assolute

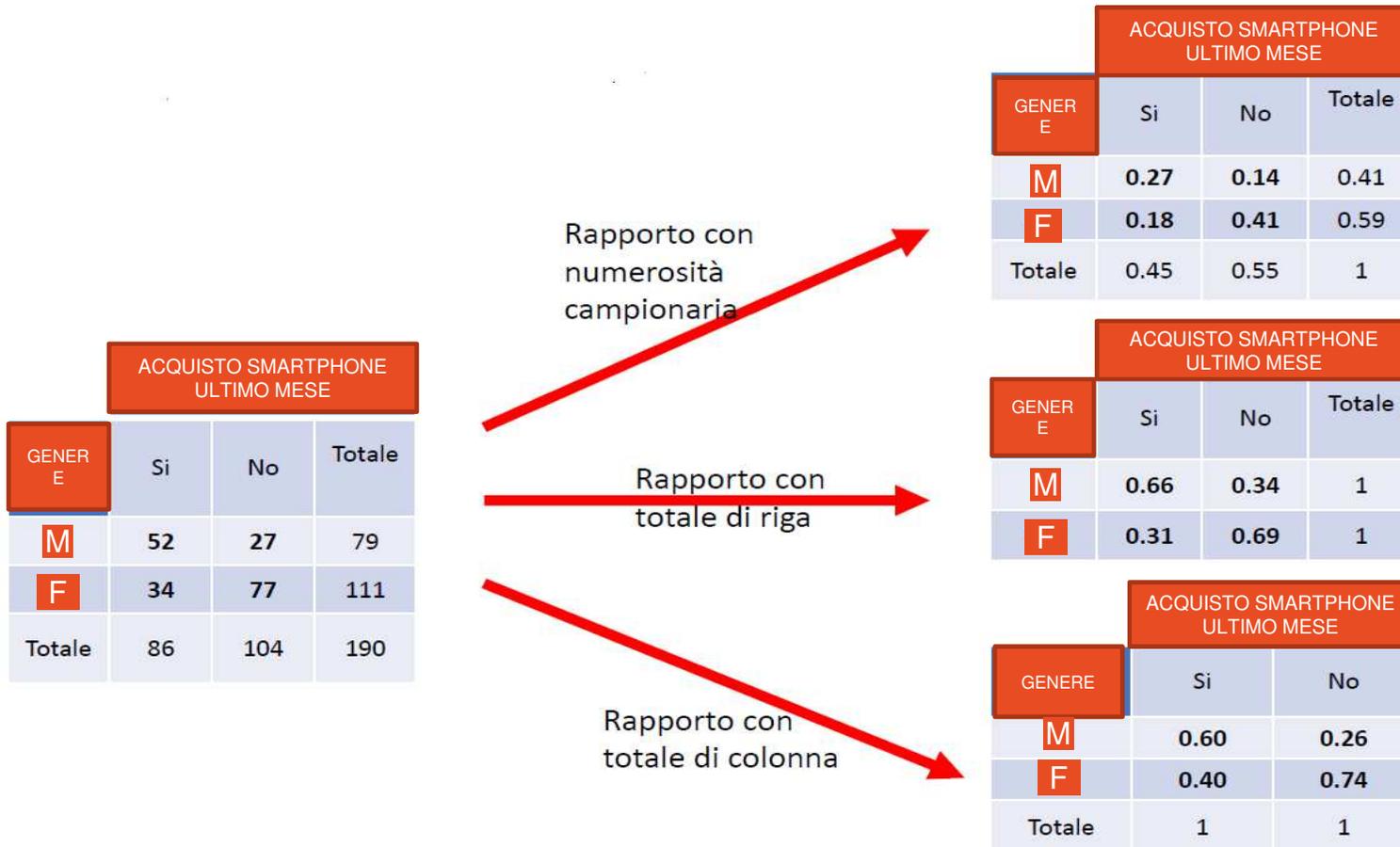
		ACQUISTO SMARTPHONE ULTIMO MESE		
GENERE		Si	No	Totale
M		52	27	79
F		34	77	111
Totale		86	104	190

27= numero dei maschi che nell'ultimo mese NON hanno acquistato uno smartphone.

TOTALE MARGINALE DI COLONNA

TOTALE MARGINALE DI RIGA

# Tablelle doppie: frequenze relative



# Distribuzioni multiple

In generale si parla di:

- **distribuzione tripla**, se si considerano congiuntamente tre caratteri (tre colonne della matrice dei dati); **distribuzione quadrupla**, se si considerano congiuntamente quattro caratteri, e così via
- In generale si parla di distribuzioni multivariate

# Serie storiche

- Si ha una serie storica quando i dati statistici di interesse vengono associati a **modalità temporali**

	2022												2023	
	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novembre	Dicembre	Gennaio	Febbraio
Clima di Fiducia dei consumatori	114.2	112.4	100.8	100.0	102.7	98.3	94.8	98.3	94.8	90.1	98.1	102.5	100.9	104.0
Clima economico	129.7	129.4	98.2	97.3	103.6	93.9	84.9	92.9	81.3	77.6	95.2	106.3	107.6	114.5
Clima personale	109.0	106.8	101.7	100.9	102.4	99.8	98.1	100.2	99.3	94.3	99.0	101.2	98.6	100.5
Clima corrente	114.7	109.6	105.7	100.8	104.6	97.9	96.1	99.7	96.9	91.0	94.9	98.6	95.7	97.6
Clima futuro	113.5	116.6	93.5	98.9	99.8	98.8	92.9	96.4	91.8	88.8	102.8	108.2	108.6	113.4

Fonte Istat: **FIDUCIA DEI CONSUMATORI** - indice (in base 2010) destagionalizzato

## Febbraio 2023:

Il clima di fiducia dei consumatori torna ad aumentare dopo la diminuzione registrata a gennaio. Il miglioramento dell'indice è dovuto ad un'evoluzione positiva delle opinioni sulla situazione economica generale, ad un aumento delle aspettative sulla situazione economica familiare e a valutazioni in miglioramento sia sull'opportunità di effettuare acquisti nella fase attuale sia sulla possibilità di risparmiare in futuro.

# Clima di fiducia dei consumatori:

È elaborato sulla base di nove domande ritenute maggiormente idonee per valutare l'ottimismo/pessimismo dei consumatori:

## Clima di Fiducia dei consumatori

- ✓ giudizi e attese sulla **situazione economica dell'Italia**;
  - ✓ attese sulla **disoccupazione**;
- ✓ giudizi e attese sulla **situazione economica della famiglia**;
  - ✓ opportunità attuale e possibilità future del **risparmio**;
    - ✓ opportunità all'**acquisto di beni durevoli**;
    - ✓ giudizi sul **bilancio familiare**

- **Clima economico:** giudizi e attese sulla situazione economica dell'Italia, attese sulla disoccupazione
- **Clima personale:** giudizi e attese sulla situazione economica della famiglia; opportunità attuale e possibilità future del risparmio; opportunità all'acquisto di beni durevoli; bilancio finanziario della famiglia
- ❖ **Clima corrente:** situazione economica dell'Italia e della famiglia; opportunità attuale del risparmio e acquisto di beni durevoli; bilancio finanziario della famiglia
- ❖ **Clima futuro:** situazione economica dell'Italia e della famiglia; disoccupazione; possibilità future di risparmio

# Serie territoriali

- Si ha una serie territoriale quando i dati vengono associati a modalità rappresentate da **entità territoriali**

ATTIVITÀ ECONOMICHE	Nord-ovest	Nord-est	Centro	Sud	Isole	Totale Italia
<b>Attività creative, artistiche e di intrattenimento</b>	<b>11086</b>	<b>5707</b>	<b>10021</b>	<b>3649</b>	<b>1487</b>	<b>31950</b>
Rappresentazioni artistiche	3709	1998	3509	1237	468	10921
Attività di supporto alle rappresentazioni artistiche	2031	1085	1921	855	335	6227
Creazioni artistiche e letterarie	5292	2580	4519	1485	669	14545
Gestione di strutture artistiche (teatri, sale da concerto e altro)	54	44	72	72	15	257
<b>Attività di biblioteche, archivi, musei ed altre attività culturali</b>	<b>233</b>	<b>197</b>	<b>258</b>	<b>183</b>	<b>160</b>	<b>1031</b>
Attività di biblioteche ed archivi	140	96	119	73	46	474
Attività di musei	34	42	77	36	38	227
Gestione di luoghi e monumenti storici e attrazioni simili	34	33	43	43	65	218
Attività degli orti botanici, dei giardini zoologici e delle riserve naturali	25	26	19	31	11	112
<b>Attività riguardanti le lotterie, le scommesse, le case da gioco</b>	<b>1901</b>	<b>1225</b>	<b>1475</b>	<b>2763</b>	<b>1001</b>	<b>8365</b>
<b>Attività sportive</b>	<b>5344</b>	<b>3651</b>	<b>3494</b>	<b>2061</b>	<b>905</b>	<b>15455</b>
<b>Attività ricreative e di divertimento</b>	<b>2990</b>	<b>2638</b>	<b>3662</b>	<b>3795</b>	<b>1382</b>	<b>14467</b>

Fonte Istat: STATISTICHE CULTURALI - ANNO 2021